



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/766,308	01/27/2004	Kas Kasravi	200901335-1	1219

22879 7590 05/27/2010

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
3404 E. Harmony Road
Mail Stop 35
FORT COLLINS, CO 80528

EXAMINER

SAINT CYR, LEONARD

ART UNIT	PAPER NUMBER
----------	--------------

2626

NOTIFICATION DATE	DELIVERY MODE
-------------------	---------------

05/27/2010

ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

JERRY.SHORMA@HP.COM
ipa.mail@hp.com
laura.m.clark@hp.com



UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

**BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES**

Application Number: 10/766,308
Filing Date: January 27, 2004
Appellant(s): KASRAVI ET AL.

Kasravi et al.
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 02/12/10 appealing from the Office action mailed 08/10/09.

(1) Real Party in Interest

A statement identifying by name the real party in interest is contained in the brief.

(2) Related Appeals and Interferences

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) Status of Claims

The statement of the status of claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Claimed Subject Matter

The summary of claimed subject matter contained in the brief is correct.

(6) Grounds of Rejection to be Reviewed on Appeal

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

(7) Claims Appendix

The copy of the appealed claims contained in the Appendix to the brief is correct.

(8) Evidence Relied Upon

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 101

1. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1, 2, 3, 6, 7, 10 - 14, 16 – 21, 34 and 36 – 40 are rejected under 35 U.S.C. 101 as not falling within one of the four statutory categories of invention. Supreme Court precedent¹ and recent Federal Circuit decisions² indicate that a statutory “process” under 35 U.S.C. 101 must (1) be tied to another statutory category (such as a particular apparatus), or (2) transform underlying subject matter (such as an article or material) to a different state or thing. While the instant claim(s) recite a series of steps or acts to be performed, the claim(s) neither transform underlying subject matter nor positively tie to another statutory category that accomplishes the claimed method steps, and therefore do not qualify as a statutory process. Claims 1, 2, 3, 6, 7, 10 - 14, 16 – 21, 34 and 36 – 40 do not identify the apparatus that accomplishes the method steps like “processing the semantic vector by a digital computer” described in

¹ *Diamond v. Diehr*, 450 U.S. 175, 184 (1981); *Parker v. Flook*, 437 U.S. 584, 588 n.9 (1978); *Gottschalk v. Benson*, 409 U.S. 63, 70 (1972); *Cochrane v. Deener*, 94 U.S. 780, 787-88 (1876).

² *In re Bilski*, 88 USPQ2d 1385 (Fed. Cir. 2008).

Art Unit: 2626

page 12, paragraph 36, last three lines of the specification. Thus claims 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 34 and 36 - 40 do not define a statutory process.

Claim Rejections - 35 USC § 102

2. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

3. Claims 1, 2, 3, 6, 7, 10 - 14, 16 - 21, 23, 24, 26 - 31, and 36 - 40 are rejected under 35 U.S.C. 102(a) as being anticipated by Gillis (US Patent 6,523,026).

As per claims 1, Gillis teaches a computer-implemented method for comparing semantic content of two or more documents, comprising:

accessing a plurality of documents ("source and target domains"); performing a linguistic analysis on each document ("computing a set of vectors"; col.10, lines 9 - 17); col.11, lines 36 - 40);

defining a semantic vector for each document based on the linguistic analysis, said semantic vector having multiple components, wherein each component of said semantic vector has at least: a term included in the document or a synonym of said term; a weighting factor relating to an importance ("Selected terms within a multiterm query may be weighted, if desired, to reflect their importance to the user"), based on characteristics of the document of said term ("**The weightings of word relationships are entirely natural domain specific**"); and a frequency value relating to a number of occurrences of said term ("computing a set of term vectors"; col.11, lines 36, and 37; col.10, lines 18 - 20; col.42, line 67- col.43, line 1; col.23, lines 35 -40); and

comparing a semantic vector of an identified document to the semantic vector for each document in the plurality of documents to determine at least one document semantically similar to the identified document (**“Semantic similarity of terms is compared by computing vector overlap (i.e., the dot products of the term vectors).** The degree of overlap of the vectors reflects the degree to which two terms have a similar pattern of usage (or “meaning”) in the training corpus” (col.7, lines 59 – 66; col. 33, line 66 – col.34, line 9; col.40, lines 1 - 5).

As per claim 2, Gillis further discloses that the linguistic analysis comprises sentence analysis (“sentence in the individual documents”; col.43, lines 43 – 46).

As per claim 3, Gillis further discloses that the sentence analysis comprises a syntactic analysis (“preferred stop list word include in the vectorization”) and a semantic analysis (“semantic similarity”; col.39, lines 14 – 20; col.35, lines 4 – 6).

As per claim 6, Gillis further discloses that each component of the semantic vector for at least one of the documents comprises multiple dimensions (“n dimensional space”; col.39, line 63 –col.40, line 1).

As per claim 7, Gillis further discloses that each component of the semantic vector for at least one of the documents further comprises a subordinate concept value (“cable” is the subordinate concept of term “telecommunications”; col.51, lines 30 – 35).

As per claim 10, Gillis further discloses that some of the components of the semantic vector have for at least one of the documents {main term – subordinate term pairs} as their first value (“cable” and “telecommunications” are related term pairs, wherein cable is the subordinate term of telecommunications; col.51, lines 30 – 35).

As per claim 11, Gillis further discloses that the semantic vector comprises a multi-dimensional vector defined by the content of a semantic net (“n dimensional semantic space”; col.39, line 63 – col.40, line 1).

As per claim 12, Gillis further discloses that the content of the semantic net is augmented by relative weights, strengths, or frequencies of occurrence of the features within the semantic net (“frequency related weightings to term in the computation of summary vectors”; col.41, lines 40 - 46).

As per claim 14, 23, 24, 26, and 27, Gillis teaches comparing two or more documents, by:

linguistically analyzing a plurality of documents to identify at least one term group in each document, each tem group comprising a main term and at least one subordinate term semantically related to the main term (“a small subset of terms (or groups of terms such as phrases) is chosen from the source domain...computing a set of vectors”; col.10, lines 9 – 22); col.11, lines 36 – 40);

generating a semantic vector associated with each document, the semantic vector comprising a plurality of components, each component including; a term group in the document; a frequency value relating to a number of occurrences of the term group; and a weighting factor relating to an importance (“Selected terms within a multiterm query may be weighted, if desired, to reflect their importance to the user”), based on characteristics of the document of said term (**“The weightings of word relationships are entirely natural domain specific”**) of at least part of the term group (“computing a set of term vectors... Vector of terms that occurred less frequently in the training corpus are weighted more heavily in the calculation of summary vectors of search domain records”; col.11, lines 36, and 37; col.10, lines 18 – 20; col.42, line 67- col.43, line 1; col.41, lines 43 – 47; col.23, lines 35 -40); and

comparing a semantic vector of an identified document to the semantic vector for each document in the plurality of documents to determine at least one document semantically similar to the identified document using a defined metric (**“Semantic similarity of terms is compared by computing vector overlap (i.e., the dot products of the term vectors)**). The degree of overlap of the vectors reflects the degree to which two terms have a similar pattern of usage (or “meaning”) in the training corpus” (col.7, lines 59 – 66; col. 33, line 66 – col.34, line 9; col.40, lines 1 - 5);

wherein said metric measures the semantic distance between documents as a function of at least the frequency values included in the semantic vectors for the documents (“semantically distant are individually represented at least 50 times”; col.48, lines 48 – 55; col.51, lines 30 – 35; col.41, lines 43 - 47).

As per claim 16, Gillis further discloses that the main term includes synonyms of the main term (col.11, line 8).

As per claims 17, 28, Gillis further discloses that one or more of said two or more documents are located using an autonomous software or 'bot program ("software programs"; col.10, lines 9 – 17; col.25, lines 57 – 67).

As per claims 18, and 29, Gillis further discloses automatically analyzes each document in a defined domain (source and target domains) or network by executing a series of rules and assigning an overall score to the document ("average of component values"; col.10, lines 9 – 17; col.41, line 66 –col.42, line 25).

As per claim 19, Gillis further discloses that all documents with a score above a defined threshold are linguistically analyzed ("generate term vectors and accept only records that match all the categories beyond some minimum threshold"; col.46, line 65 – col.47, line 11).

As per claims 20, and 30, Gillis further discloses that the semantic vector is a quantification of the semantic content of each document ("semantic vectors"; col.39, lines 14 – 20; col.1, lines 15 - 20).

As per claim 21, Gillis further discloses that each component has multiple dimensions (“n dimensional semantic space”; col.39, line 63 – col.40, line 1).

As per claim 31, Gillis further discloses that the output of said defined algorithm is a measure of at least one of semantic distance, semantic similarity, semantic dissimilarity, degree of patentable novelty and degree of anticipation (“semantic similarity”; col.4, lines 1 – 3).

As per claim 36, Gillis further discloses that said term comprises at least one of a word or a phrase (“a small subset of terms (or groups of terms such as phrases) is chosen from the source domain”; col.10, lines 19 – 22).

As per claim 37, Gillis further discloses that comparing the semantic vectors based on a defined algorithm (col.42, line 2).

As per claim 13, Gillis further discloses that the output of said defined algorithm is a measure of at least one of semantic distance, semantic similarity, semantic dissimilarity, degree of patentable novelty and degree of anticipation (“semantic similarity”; col.4, lines 1 – 3; col. 33, line 66 – col.34, line 9; col.40, lines 1 - 5).

As per claim 38, Gillis further discloses that the at least one subordinate term includes synonyms of one of the subordinate terms (col.11, line 8).

As per claim 39, Gillis further discloses that one or more of the at least one subordinate term or the main term comprises a phrase (col.10, lines 19 – 22).

As per claim 40, Gillis further discloses that the weighting factor comprises a plurality of different weighting factors and each of the different weighting factors relates to the importance of the main term or a subordinate term in the term group (“Vector of terms that occurred less frequently in the training corpus are weighted more heavily in the calculation of summary vectors of search domain records”; col.41, lines 43 - 47).

Allowable Subject Matter

4. Claims 33, and 35 are allowed over the prior art. The following is an examiner’s statement of reasons for allowance:

As to claim 33, and 35, Gillis does not teach or suggest that the defined metric is one of: $\text{Sqrt} (f_1^2 + f_2^2 + f_3^2 + f_4^2 + \dots + f_{(N-1)}^2 + f_N^2) / n * 100$, wherein f is a difference in frequency of a common term between two documents and n is the number of terms those documents have in common; or $\text{Sqrt}(\text{sum}((w-\text{Delta})^2 * w-\text{Avg})) / (\text{Log}(n)^3 * 1000)$, wherein $w-\text{Delta}$ is the difference in weight between two common terms, $w-\text{Avg}$ is the average weight between two common terms, and n is the number of common terms, between two documents.

(10) Response to Argument

Appellants argue that Gillis does not teach comparing a semantic vector of an identified document to the semantic vector for each document in the plurality of documents to determine at least one document semantically similar to the identified document (Appeal Brief, pages 8, and 9).

The examiner disagrees, since Gillis discloses that "relative meanings of terms in the context of the source domain can be quantitatively **compared simply by calculating the dot products (inner products) of their vectors**. Geometrically this is equivalent to comparing the relative alignment **of the vectors in the high dimensionality space ["semantic" space]**. **A high overlap (corresponding to a high dot product) indicates a close similarity of meaning**. On the other hand, terms having dissimilar meanings have **term vectors which are orthogonal (or nearly orthogonal) in the semantic space (i.e., have dot products of zero or close to zero)**. **Semantic similarity of terms is compared by computing vector overlap (i.e., the dot products of the term vectors)**. The degree of overlap of the vectors reflects the degree to which two terms have a similar pattern of usage (or "meaning") in the training corpus" (col. 33, line 66 – col.34, line 9; col.40, lines 1 - 5).

Appellants argue that the cited teaching of the Gillis patent is taken out of context, in that it describes a difficulty the disclosure of the Gillis Patent seeks to overcome. Accordingly, this cited feature is not to be included with the features of the Gillis patent cited against the claims (Appeal Brief, page 9).

Art Unit: 2626

The examiner disagrees, and points out that the cited portion (col.7, lines 59 – 66), which is similar to portions (col. 33, line 66 – col.34, line 9; col.40, lines 1 – 5) in the detailed description, was used to teach semantic vectors comparison, but not latent semantic indexing or LSI as argued by the appellants. Thus, the feature of the cited portion does not teach against the claims.

(11) Related Proceeding(s) Appendix

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

/Leonard Saint-Cyr/

Examiner, Art Unit 2626

/Richemond Dorvil/

Supervisory Patent Examiner, Art Unit 2626

Conferees:

Richemond Dorvil

/R. D./

Supervisory Patent Examiner, Art Unit 2626

Talivaldis Smits

Application/Control Number: 10/766,308
Art Unit: 2626

Page 13

Talivaldis Smits